

## PATENT ABSTRACTS OF JAPAN

(11)Publication number : 06-004446

(43)Date of publication of application : 14.01.1994

(51)Int.Cl.

G06F 13/10

G06F 9/46

(21)Application number : 04-314685

(71)Applicant : FUJITSU LTD

(22)Date of filing : 25.11.1992

(72)Inventor : UCHINO MINORU

(30)Priority

Priority number : 04101495

Priority date : 22.04.1992

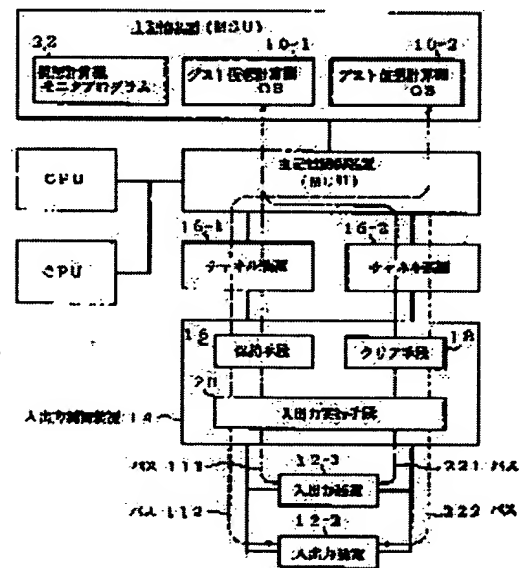
Priority country : JP

## (54) METHOD AND DEVICE FOR INPUT/OUTPUT CONTROL OF VIRTUAL COMPUTER SYSTEM

(57)Abstract:

PURPOSE: To improve reliability and performance, and to reduce the overhead of a VM monitor in the case of executing input/output control by allocating a specific device to plural guest VMs(virtual computers).

CONSTITUTION: Every plural access paths for input/output devices 12-1, 12-2 allocated to guest VMs 10-1, 10-2 one-to-one, a table in which a group ID issued from each of the guest VMs 10-1, 10-2 is stored is held in a holding means 16, and when an input/output instruction issued from the guest VMs 10-1, 10-2 is received, the access path belonging to the same group is detected from among plural access paths for the input/output device by referring to the table of the holding means 16 with respect to a group to which the input/output device executing the input/output instruction belongs, and by using one of the access paths detected by an input/output executing means 20, an input and an output between the guest VMs 10-1, 10-2 are executed. Even in a state that a single magnetic disk device is shared between plural guest VMs, exclusive control is executed without necessitating an intervention of a VM monitor.



## LEGAL STATUS

[Date of request for examination]

25.11.1992

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

(19) 日本国特許庁 (J P)

(12) 特 許 公 報 (B 2)

(11) 特許出願公告番号

特公平7-97358

(24) (44) 公告日 平成 7 年 (1995) 10 月 18 日

(51) Int.Cl. <sup>6</sup>	識別記号	庁内整理番号	F I	技術表示箇所
G 0 6 F 13/10	3 3 0 C	7368-5B		
9/46	3 5 0	7737-5B		

請求項の数28(全 20 頁)

(21) 出願番号	特願平4-314685	(71) 出願人	000005223 富士通株式会社 神奈川県川崎市中原区上小田中1015番地
(22) 出願日	平成 4 年 (1992) 11 月 25 日	(72) 発明者	内野 実 神奈川県川崎市中原区上小田中1015番地 富士通株式会社内
(65) 公開番号	特開平6-4446	(74) 代理人	弁理士 竹内 進 (外 1 名)
(43) 公開日	平成 6 年 (1994) 1 月 14 日		
(31) 優先権主張番号	特願平4-101495		
(32) 優先日	平 4 (1992) 4 月 22 日		
(33) 優先権主張国	日本 (J P)		

審査官 大日方 和幸

(54) 【発明の名称】 仮想計算機システムの入出力制御方法及び装置

【 特許請求の範囲】

【 請求項1 】 仮想計算機システムを制御する仮想計算機モニタプログラム ( 2 2 ) で実現される複数のゲスト仮想計算機 ( 1 0 - 1 , 1 0 - 2 ) の各々に、複数の入出力装置 ( 1 2 - 1 , 1 2 - 2 ) の中のいずれか 1 つを 1 対 1 に割当てる割当過程と；該割当過程の結果に基づき、前記複数のゲスト仮想計算機 ( 1 0 - 1 , 1 0 - 2 ) 、複数のチャネル装置 ( 1 6 - 1 , 1 6 - 2 ) 及び入出力制御装置 ( 1 4 ) を経由して前記入出力装置 ( 1 2 - 1 , 1 2 - 2 ) に至る複数のアクセスパス ( 1 1 1 , 1 1 2 , 2 2 1 , 2 2 2 ) 毎に、前記ゲスト仮想計算機 ( 1 0 - 1 , 1 0 - 2 ) の各々が設定したグループ I D を入出力制御装置 ( 1 4 ) 内のテーブルに保持するグループ I D 保持過程と；前記仮想計算機 ( 1 0 - 1 , 1 0 - 2 ) のいずれかで発行した入出力命令を前記入

出力制御装置 ( 1 4 ) で受付けた際に、前記テーブルを参照して命令で指定された入出力装置に至る複数のアクセスパスの中から同一グループに属するアクセスパスを検出し、入出力を実行する入出力実行過程と；を備えたことを特徴とする仮想計算機システムの入出力制御方法。

【 請求項2 】 請求項1 記載の仮想計算機システムの入出力制御方法に於いて、前記グループ I D 保持過程は、仮想計算機モニタプログラム ( 2 2 ) による入出力装置 ( 1 2 - 1 , 1 2 - 2 ) の割当て結果に基づいて前記ゲスト仮想計算機 ( 1 0 - 1 , 1 0 - 2 ) を起動し、該仮想計算機上で動作するオペレーティングシステム ( O S ) のグループ I D 設定処理によりグループ I D の設定を受けることを特徴とする仮想計算機システムの入出力制御方法。

【 請求項3 】 請求項1 記載の仮想計算機システムの入出

力制御方法に於いて、更に、仮想計算機モニタプログラム( 2 2 ) の指示により前記入出力制御装置( 1 4 ) に設けたクリア手段( 1 8 ) が特定のグループI Dをクリアするクリア過程を設けたことを特徴とする仮想計算機システムの入出力制御方法。

【請求項4】請求項3 記載の仮想計算機システムの入出力制御方法に於いて、前記クリア過程は、前記仮想計算機モニタプログラム( 2 2 ) によるゲスト 仮想計算機( 1 0 - 1 , 1 0 - 2 ) に対する入出力装置( 1 2 - 1 , 1 2 - 2 ) の割当て変更時に、前記入出力制御装置( 1 4 ) のクリア手段( 1 8 ) にクリアを指示して該ゲスト 仮想計算機( 1 0 - 1 , 1 0 - 2 ) が使用していたグループI Dをクリアさせることを特徴とする仮想計算機システムの入出力制御方法。

【請求項5】請求項1 記載の仮想計算機システムの入出力制御方法に於いて、前記グループI D保持過程は、前記入出力制御装置( 1 4 ) に、前記複数のチャネル装置( 1 6 - 1 , 1 6 - 2 ) ごとに前記グループI Dを複数設定するテーブルを設けたことを特徴とする仮想計算機システムの入出力制御方法。

【請求項6】請求項1 記載の仮想計算機システムの入出力制御方法に於いて、入出力実行過程は、前記テーブルを参照して検出された同一のグループに属するアクセスパスを使用して入出力処理を実行することを特徴とする仮想計算機システムの入出力制御方法。

【請求項7】請求項1 乃至6 記載の仮想計算機システムの入出力制御方法に於いて、前記入出力装置( 1 2 - 1 , 1 2 - 2 ) は磁気ディスク装置であり、入出力制御装置( 1 4 ) はディスク制御装置であることを特徴とする仮想計算機システムの入出力制御方法。

【請求項8】仮想計算機システムを制御する仮想計算機モニタプログラム( 2 2 ) で実現される複数のゲスト 仮想計算機( 1 0 - 1 , 1 0 - 2 ) と; 該複数のゲスト 仮想計算機( 1 0 - 1 , 1 0 - 2 ) の各々に1 対1 に割当てられた複数の入出力装置( 1 2 - 1 , 1 2 - 2 ) と、該複数の入出力装置( 1 2 - 1 , 1 2 - 2 ) の各々に対して少なくとも2 系統のパスで結合した入出力制御装置( 1 4 ) と、

該入出力制御装置( 1 4 ) を前記複数のゲスト 仮想計算機( 1 0 - 1 , 1 0 - 2 ) に結合する少なくとも2 つのチャネル装置( 1 6 - 1 , 1 6 - 2 ) とを備え、

前記入出力制御装置( 1 4 ) に、前記ゲスト 仮想計算機( 1 0 - 1 , 1 0 - 2 ) からチャネル装置( 1 6 - 1 , 1 6 - 2 ) 及び入出力制御装置( 1 4 ) を経由して前記入出力装置( 1 2 - 1 , 1 2 - 2 ) に至る複数のアクセスパス( 1 1 1 , 1 1 2 , 2 2 1 , 2 2 2 ) 毎に、前記ゲスト 仮想計算機( 1 0 - 1 , 1 0 - 2 ) の各々が設定したグループI Dをテーブルに保持するグループI D保持手段( 1 6 ) と、

該グループI D保持手段( 1 6 ) のテーブルに保持した

特定のグループI Dをクリアするクリア手段( 1 8 )

と、

前記仮想計算機( 1 0 - 1 , 1 0 - 2 ) のいずれかが発行した入出力命令を受付けた際に、前記グループI D保持手段( 1 6 ) のテーブルを参照して命令で指定された入出力装置に至る複数のアクセスパスの中から同一グループに属するアクセスパスを検出し、入出力を実行する入出力実行手段( 2 0 ) と、を設けたことを特徴とする仮想計算機システムの入出力制御装置。

【請求項9】請求項8 記載の仮想計算機システムの入出力制御装置に於いて、前記グループI D保持手段( 1 6 ) は、仮想計算機モニタプログラム( 2 2 ) による入出力装置( 1 2 - 1 , 1 2 - 2 ) の割当て結果に基づいて前記ゲスト 仮想計算機( 1 0 - 1 , 1 0 - 2 ) を起動し、該仮想計算機上で動作するオペレーティングシステム( OS ) のグループI D設定処理によるグループI Dの設定を受けることを特徴とする仮想計算機システムの入出力制御装置。

【請求項1 0】請求項8 記載の仮想計算機システムの入出力制御装置に於いて、前記入出力制御装置( 1 4 ) に設けたクリア手段( 1 8 ) は、仮想計算機モニタプログラム( 2 2 ) の指示により特定のグループI Dをクリアすることを特徴とする仮想計算機システムの入出力制御装置。

【請求項1 1】請求項1 0 記載の仮想計算機システムの入出力制御装置に於いて、前記仮想計算機モニタプログラム( 2 2 ) は、入出力装置( 1 2 - 1 , 1 2 - 2 ) のゲスト 仮想計算機( 1 0 - 1 , 1 0 - 2 ) への割り当て変更時に、該ゲスト 仮想計算機( 1 0 - 1 , 1 0 - 2 ) が使用していたグループI Dをクリアするために前記入出力制御装置( 1 4 ) のクリア手段( 1 8 ) にクリアを指示することを特徴とする仮想計算機システムの入出力制御装置。

【請求項1 2】請求項8 記載の仮想計算機システムの入出力制御装置に於いて、前記入出力制御装置( 1 4 ) は前記複数のチャネル装置( 1 6 - 1 , 1 6 - 2 ) ごとに前記グループI Dを複数設定するテーブルを設けたことを特徴とする仮想計算機システムの入出力制御装置。

【請求項1 3】請求項8 記載の仮想計算機システムの入出力制御装置に於いて、前記入出力実行手段( 2 0 ) は、前記保持手段( 1 6 ) のテーブルから得られたグループI Dで指定された2 つ以上のアクセスパスを使用して入出力処理を実行することを特徴とする仮想計算機システムの入出力制御装置。

【請求項1 4】請求項8 乃至1 3 記載の仮想計算機システムの入出力制御装置に於いて、前記入出力装置( 1 2 - 1 , 1 2 - 2 ) は磁気ディスク装置であり、入出力制御装置( 1 4 ) はディスク制御装置であることを特徴とする仮想計算機システムの入出力制御装置。

【請求項1 5】仮想計算機システムを制御する仮想計算

機モニタプログラム( 2 2 ) で実現される複数のゲスト仮想計算機( 1 0 - 1 , 1 0 - 2 ) に、チャンネル装置( 1 6 - 1 )、入出力制御装置( 1 4 ) を介して少なくとも1 つの物理アクセスパス( 1 1 1 ) と2 つの論理アクセスパス( 1 1 1 1 , 2 1 1 1 ) で結合される1 つの入出力装置( 1 2 - 1 ) を共通に割当てて割り当て過程と; 該割り当て過程の結果に基づき、前記論理アクセスパス( 1 1 1 1 , 2 1 1 1 ) と前記物理アクセスパス( 1 1 1 ) との対応関係、および前記論理アクセスパス( 1 1 1 1 , 2 1 1 1 ) ごとに前記ゲスト仮想計算機( 1 0 - 1 , 1 0 - 2 ) の各々が設定したグループI D を入出力制御装置( 1 4 ) 内のテーブルに保持するグループI D 保持過程と; 前記仮想計算機( 1 0 - 1 , 1 0 - 2 ) のいずれかで発行した入出力命令を前記入出力制御装置( 1 4 ) で受付けた際に、前記テーブルを参照して論理アクセスパスを物理アクセスパスに変換して、物理アクセスパスの競合を制御する排他制御過程と; 該排他制御過程の制御結果に基づいて前記入出力装置( 1 2 - 1 ) に対する入出力命令を実行する入出力実行過程と; を備えたことを特徴とする仮想計算機システムの入出力制御方法。

【請求項1 6】請求項1 5 記載の仮想計算機システムの入出力制御方法に於いて、前記グループI D 保持過程は、仮想計算機モニタプログラム( 2 2 ) による入出力装置( 1 2 - 1 ) の割当て結果に基づいて前記ゲスト仮想計算機( 1 0 - 1 , 1 0 - 2 ) を起動し、該仮想計算機上で動作するオペレーティングシステム( OS ) のグループI D 設定処理によりグループI D の設定を受けることを特徴とする仮想計算機システムの入出力制御方法。

【請求項1 7】請求項1 5 記載の仮想計算機システムの入出力制御方法に於いて、更に、仮想計算機モニタプログラム( 2 2 ) の指示により前記入出力制御装置( 1 4 ) に設けたクリア手段( 1 8 ) が特定のグループI D をクリアするクリア過程を設けたことを特徴とする仮想計算機システムの入出力制御方法。

【請求項1 8】請求項1 7 記載の仮想計算機システムの入出力制御方法に於いて、前記クリア過程は、前記仮想計算機モニタプログラム( 2 2 ) によるゲスト仮想計算機( 1 0 - 1 ) に対する入出力装置( 1 2 - 1 ) の割り当て変更時に、前記入出力制御装置( 1 4 ) のクリア手段( 1 8 ) にクリアを指示して該ゲスト仮想計算機( 1 0 - 1 , 1 0 - 2 ) が使用していたグループI D をクリアさせることを特徴とする仮想計算機システムの入出力制御方法。

【請求項1 9】請求項1 5 記載の仮想計算機システムの入出力制御方法に於いて、前記グループI D 保持過程は、前記複数のチャンネル装置( 1 6 - 1 , 1 6 - 2 ) ごとに前記グループI D を複数設定するテーブルを設けたことを特徴とする仮想計算機システムの入出力制御方

法。

【請求項2 0】請求項1 5 記載の仮想計算機システムの入出力制御方法に於いて、前記ゲスト仮想計算機( 1 0 - 1 , 1 0 - 2 ) の各々に異なる複数の入出力装置( 1 2 - 1 , 1 2 - 2 ) を共通に割当てた場合、前記入出力実行過程は、前記テーブルから得られたグループI D で指定された1 又は複数の物理アクセスパスを使用して入出力処理を実行することを特徴とする仮想計算機システムの入出力制御方法。

10 【請求項2 1】請求項1 5 乃至2 0 記載の仮想計算機システムの入出力制御方法に於いて、前記入出力装置( 1 2 - 1 , 1 2 - 2 ) は磁気ディスク装置であり、入出力制御装置( 1 4 ) はディスク制御装置であることを特徴とする仮想計算機システムの入出力制御方法。

【請求項2 2】仮想計算機システムを制御する仮想計算機モニタプログラム( 2 2 ) で実現される複数のゲスト仮想計算機( 1 0 - 1 , 1 0 - 2 ) と; 該複数のゲスト仮想計算機( 1 0 - 1 , 1 0 - 2 ) に共通に割当てられた1 つの入出力装置( 1 2 - 1 ) と、

20 該入出力装置( 1 2 - 1 ) に1 つの物理アクセスパス( 1 1 1 ) で結合した入出力制御装置( 1 4 ) と、該入出力制御装置( 1 4 ) を介して複数の論理アクセスパス( 1 1 1 1 , 2 1 1 1 ) により前記複数のゲスト仮想計算機( 1 0 - 1 , 1 0 - 2 ) に結合する少なくとも1 つのチャンネル装置( 1 6 - 1 ) とを備え、

前記入出力制御装置( 1 4 ) に、

前記ゲスト仮想計算機( 1 0 - 1 , 1 0 - 2 ) からの入出力命令で指示された論理アクセスパス( 1 1 1 1 , 2 1 1 1 ) を物理アクセスパス( 1 1 1 ) に変換すると共に、

30 前記ゲスト仮想計算機( 1 0 - 1 , 1 0 - 2 ) から前記入出力装置( 1 2 - 1 , 1 2 - 2 ) に至る複数の論理アクセスパス( 1 1 1 1 , 2 1 1 1 ) 毎に、前記ゲスト仮想計算機( 1 0 - 1 , 1 0 - 2 ) の各々が設定したグループI D をテーブルに保持するグループI D 保持手段( 1 6 ) と、

該グループI D 保持手段( 1 6 ) のテーブルに保持した特定のグループI D をクリアするクリア手段( 1 8 ) と、

と、

40 前記ゲスト仮想計算機( 1 0 - 1 , 1 0 - 2 ) のいずれかが発行した入出力命令を受付けた際に、前記グループI D 保持手段( 1 6 ) のテーブル参照により論理アクセスパスから物理アクセスパス( 1 1 1 ) に変換してアクセスの競合を制御する排他制御手段( 4 0 ) と、

該排他制御手段( 4 0 ) の制御結果に基づいて前記入出力装置( 1 2 - 1 ) に対する入出力命令を実行する入出力実行手段( 2 0 ) と、を設けたことを特徴とする仮想計算機システムの入出力制御装置。

【請求項2 3】請求項2 2 記載の仮想計算機システムの入出力制御装置に於いて、前記入出力制御装置( 1 4 ) に設けたグループI D 保持手段( 1 6 ) は、仮想計算機

5

モニタプログラム( 2 2 ) による特定の入出力装置( 1 2 - 1 ) の割当て結果に基づいて前記ゲスト 仮想計算機( 1 0 - 1 , 1 0 - 2 ) を起動し、該仮想計算機上で動作するオペレーティングシステム( OS ) のグループI D 設定処理によりグループI D の設定を受けることを特徴とする仮想計算機システムの入出力制御装置。

【請求項2 4】請求項2 2 記載の仮想計算機システムの入出力制御装置に於いて、前記入出力制御装置( 1 4 ) に設けたクリア手段( 1 8 ) は、仮想計算機モニタプログラム( 2 2 ) の指示により特定のグループI D をクリアすることを特徴とする仮想計算機システムの入出力制御装置。

【請求項2 5】請求項2 4 記載の仮想計算機システムの入出力制御装置に於いて、前記仮想計算機モニタプログラム( 2 2 ) は、入出力装置( 1 2 - 1 ) のゲスト 仮想計算機( 1 0 - 1 , 1 0 - 2 ) への割り当て変更時に、該ゲスト 仮想計算機( 1 0 - 1 , 1 0 - 2 ) が使用していたグループI D をクリアするために前記入出力制御装置( 1 4 ) のクリア手段( 1 8 ) にクリアを指示することを特徴とする仮想計算機システムの入出力制御装置。

【請求項2 6】請求項2 2 記載の仮想計算機システムの入出力制御装置に於いて、前記入出力制御装置( 1 4 ) は前記複数のチャネル装置( 1 6 - 1 , 1 6 - 2 ) ごとに前記グループI D を複数設定するテーブルを設けたことを特徴とする仮想計算機システムの入出力制御装置。

【請求項2 7】請求項2 2 記載の仮想計算機システムの入出力制御装置に於いて、前記入出力実行手段( 2 0 ) は、前記ゲスト 仮想計算機( 1 0 - 1 , 1 0 - 2 ) の各々に複数の入出力装置( 1 2 - 1 , 1 2 - 2 ) を割当てた場合、前記グループI D 保持手段( 1 6 ) のテーブルから得られた同じグループI D をもつ2 つ以上の物理アクセスパスを使用して入出力処理を実行することを特徴とする仮想計算機システムの入出力制御装置。

【請求項2 8】請求項2 2 乃至2 7 記載の仮想計算機システムの入出力制御装置に於いて、前記入出力装置( 1 2 - 1 , 1 2 - 2 ) は磁気ディスク装置であり、入出力制御装置( 1 4 ) はディスク制御装置であることを特徴とする仮想計算機システムの入出力制御装置。

【発明の詳細な説明】

【0 0 0 1】

【産業上の利用分野】本発明は、仮想計算機システム上で動作する複数のオペレーティングシステム( 以下「OS」という) で実現される複数のゲスト 仮想計算機( 以下「ゲスト VM」という) に複数の磁気ディスク装置等の入出力装置の中の特定の装置を割当てて入出力制御を行う仮想計算機システムの入出力制御方法および装置に関する。

【0 0 0 2】複数のOS で実現されるゲスト VM でディスク制御装置を共用し、ディスク制御装置の配下にある複数の磁気ディスク装置を特定のゲスト VM に1 対1 に

6

割当てて入出力を行う場合、仮想計算機システムのオーバーヘッドが増加し性能が悪化する。一方、ディスク制御装置の配下の1 つのディスク装置を複数のゲスト VM で共有する場合、物理アクセスパスを各ゲスト VM 専用に割当てると、物理アクセスパスの本数により共有するゲスト VM のが制限される。そこで、1 本の物理アクセスパスを複数のゲスト VM に割り当てる方法がある。

【0 0 0 3】しかし、1 本の物理アクセスパスを複数のゲスト VM に割り当てると、ディスク制御装置と物理アクセスパスを複数のゲスト VM に共有させるための入出力制御が必要となり、仮想計算機システムのオーバーヘッドが増加し性能が悪化する。更にディスク制御装置は本来1 つのOS による入出力制御を予定しているため、複数のゲスト VM による制御を受けた場合には、性能向上及び信頼性向上のための多重アクセスパスの機能が利用できず、この点の改善が望まれる。

【0 0 0 4】

【従来の技術】従来、仮想計算機システムにあつては、主記憶装置( MSU ) の主記憶制御装置( MCU ) にチャネル装置を介してディスク制御装置を接続しており、ディスク制御装置の配下には複数の磁気ディスク装置が接続されている。この場合、1 台のチャネル装置には複数のディスク制御装置が接続可能である。また1 台のディスク制御装置には複数の磁気ディスク装置が接続可能である。

[ ゲスト VM に入出力装置を1 対1 に割当てた形態] 図1 4 は従来の仮想計算機システムにおける入出力サブシステムの一例を示したもので、1 台のディスク制御装置1 4 の配下に複数の磁気ディスク装置1 2 - 1 , 1 2 - 2 を2 つの経路で接続した場合の例である。

【0 0 0 5】主記憶装置( MSU ) のゲスト VM 1 0 - 1 , 1 0 - 2 に対応して2 つのチャネル装置1 6 - 1 , 1 6 - 2 が設けられ、チャネル装置1 6 - 1 , 1 6 - 2 はディスク制御装置1 4 のチャネルアダプタ2 4 - 1 , 2 4 - 2 に接続されている。磁気ディスク装置1 2 - 1 又は1 2 - 2 をアクセスするには必ずチャネル装置1 6 - 1 又は1 6 - 2 とディスク制御装置1 4 を経由することとなり、磁気ディスク装置1 2 - 1 , 1 2 - 2 をアクセスする経路をアクセスパスと呼ぶことにする。またアクセスパスを各装置に付加されている番号を併せて表現する。

【0 0 0 6】例えば、チャネル装置1 6 - 1 , 1 6 - 2 の番号をCH1 , CH2、ディスク制御装置1 4 のチャネルアダプタ番号をCA1 , CA2、磁気ディスク装置1 2 - 1 , 1 2 - 2 の番号をDEV - 1 , DEV - 2 とすると、アクセスパス1 1 1 , 1 1 2 , 2 2 1 , 2 2 2 で表現され、次の経路を意味する。

[ アクセスパス]

1 1 1 CH1 ~ CA1 ~ DEV - 1

1 1 2 CH1 ~ CA1 ~ DEV - 2

50

2 2 1 CH2 ~ CA2 ~ DE V - 1

2 2 2 CH2 ~ CA2 ~ DE V - 2

図1 4 に示すような入出力構成を使用して仮想計算機システムの運用を行う場合、ディスク制御装置1 4 配下の個々の磁気ディスク装置1 2 - 1, 1 2 - 2 を、各ゲスト VM1 0 - 1, 1 0 - 2 に1 対1 に割り当てて使用する運用形態がある。

【 0 0 0 7 】 図1 4 では磁気ディスク装置1 2 - 1 をゲスト VM1 0 - 1 に割り当てているため、アクセスパス1 1 1 とアクセスパス2 2 1 がゲスト VM1 0 - 1 で使用可能である。また磁気ディスク装置1 2 - 2 をゲスト VM1 0 - 2 に割り当てているため、アクセスパス1 1 2 とアクセスパス2 2 2 がゲスト VM1 0 - 2 で使用可能である。

【 0 0 0 8 】 このようなゲスト VM とディスク装置を1 対1 に割り当てた形態の場合、ディスク制御装置1 4 とチャネル装置1 6 - 1, 1 6 - 2 は複数のゲスト VM1 0 - 1, 1 0 - 2 間で共用されることとなり、資源の有効利用の面から効果がある。ところでディスク制御装置1 4 には、性能改善および信頼性の向上を目的とする次の機能を装備するものがある。

【 0 0 0 9 】 ①動的経路再結合機能

磁気ディスク装置1 2 - 1, 1 2 - 2 のシーク完了後の再結合等を空いているアクセスパスから行うことで入出力応答時間の短縮が可能となり、性能が向上する。

②拡張リザーブ機能

OS が磁気ディスク装置1 2 - 1 又は1 2 - 2 をリザーブ中に障害が発生した時、別のチャネルを使用して磁気ディスク装置1 2 - 1 間は1 2 - 2 へのアクセスができ、信頼性が向上する。

【 0 0 1 0 】 これらの機能を利用するためには、磁気ディスク装置1 2 - 1, 1 2 - 2 に対する複数のアクセスパスをグループ形成することが必要となる。即ち、ディスク制御装置1 4 を経由する複数のアクセスパスを1 システム ( 1 つの OS ) でのみ使用することをグループ形成という形でディスク制御装置1 4 に宣言すればよい。

【 0 0 1 1 】 例えばゲスト VM1 0 - 1 が動的経路再結合機能や拡張リザーブ機能を利用する場合には、ゲスト VM1 0 - 1 上で動作するオペレーティングシステムが、当該ゲスト VM に割当てられた磁気ディスク装置1 2 - 1 の2 つのアクセスパス1 1 1, 2 2 1 毎にグループI D 設定命令を発行し、ディスク制御装置1 4 のチャネルアダプタ2 4 - 1, 2 4 - 2 に同じグループI D を保持させて磁気ディスク装置1 2 - 1 に対する2 つのアクセスパス1 1 1, 2 2 1 のグループを形成する。

[ 複数のゲスト VM に単一の入出力装置を割当てた共用形態 ] 図1 5 は、チャネル装置1 6 - 1、ディスク制御装置1 4 および磁気ディスク装置1 2 - 1 を2 つのゲスト VM1 0 - 1, 1 0 - 2 に割当てて共通に使用する運用形態を示した説明図である。

【 0 0 1 2 】 この場合には、磁気ディスク装置1 2 - 1 に対するアクセスパス1 1 1 をゲスト VM1 0 - 1 とゲスト VM1 0 - 2 で使用することとなり、1 つのアクセスパスが複数のゲスト VM 間で共用できることから、資源の有効利用の面から効果がある。この図1 5 に示す運用形態についても、図1 4 の場合と同様に、ディスク制御装置1 4 には、性能改善および信頼性の向上を目的として、動的経路再結合機能と拡張リザーブ機能が設けられる。

【 0 0 1 3 】

【 発明が解決しようとする課題 】 しかしながら、図1 4 に示したゲスト VM に磁気ディスク装置を1 対1 に割り当てて複数のゲスト VM により1 つのディスク制御装置を共用する入出力の形態にあっては、アクセスパスのグループ形成を必要とする動的経路再結合機能および拡張リザーブ機能を使用することができないという問題があった。

【 0 0 1 4 】 その理由は、本来、ディスク制御装置はチャネル装置からのチャネル・コマンド・ワード ( C C W ) による指示が1 つのシステム ( 1 つの OS ) からのものであることを前提にして動作しているためであり、図1 4 のような複数のゲスト VM1 0 - 1, 1 0 - 2 からの C C W の区別ができないことに起因している。そのため従来の仮想計算機システムにあっては、VM モニタは、システム起動時にグループI D 設定命令が複数のゲスト VM から発行され、その後の入出力動作が保証できなくなることを防止するために、ゲスト VM が発行する入出力命令を全てインタセプトし、入出力要求の中のグループI D 設定命令をエラーとしてシミュレーションする方法を採用し、グループ形成コマンドをリジェクトしている。

【 0 0 1 5 】 この入出力要求の中のグループI D 設定命令をエラーとするシミュレーションは、ゲスト VM が入出力命令を発行したら、VM モニタがグループI D 設定命令であるかを監視し、グループI D 設定命令を検知するとコマンドリジェクトを行う。その結果、複数のゲスト VM で1 つのディスク制御装置を共用する形態では、アクセスパスのグループ形成を必要とする動的経路再結合機能および拡張リザーブ機能が使用できず、性能悪化と信頼性低下を招いている。

【 0 0 1 6 】 更に、VM モニタは、エラーシミュレーションによってグループI D 設定命令をリジェクトするI O 命令への介入を実施しているため、VM モニタのオーバーヘッドの増加となり、性能悪化を助長するという問題があった。一方、図1 5 に示した1 つの磁気ディスク装置を複数のゲスト VM で共用する入出力の形態では、ゲスト VM1 0 - 1, 1 0 - 2 から同時に入出力要求が行われる場合もあることから、ディスク制御装置1 4 に排他制御を行う機能を装備することになる。

【 0 0 1 7 】 しかし、ディスク制御装置1 4 は、チャネ

ル装置16-1からのチャネル・コマンド・ワード(CCW)による指示が1つのシステム(1つのOS)からのものであることを前提に動作しているため、1つのアクセスパスを複数のシステム(複数のOS)で共用すると、排他制御ができない。そこで従来システムでは、VMモニタがディスク制御装置14に代って排他制御命令をエラーシミュレーションするために入出力命令と入出力割込みをインタセプトする必要がある、VMモニタのオーバーヘッドの増加となり、性能悪化を助長するという問題があった。

【0018】また図14の運用形態と同様に、複数のゲストVMで1つのディスク制御装置を共用する形態であることから、アクセスパスのグループ形成を必要とする動的経路再結合機能および拡張リザーブ機能が使用できず、性能悪化と信頼性低下を招いている。本発明の目的は、入出力制御装置を共用する仮想計算機システムの運用形態で、入出力制御装置が複数の仮想計算機を認識ができるようにした仮想計算機システムの入出力制御方法及び装置を提供する。

【0019】本発明の他の目的は、仮想計算機に1対1に入出力装置を割当てた場合に入出力制御装置が複数の仮想計算機を認識できるようにした仮想計算機システムの入出力制御方法及び装置を提供する。本発明の他の目的は、入出力制御装置で仮想計算機を認識してアクセスパスのグループ形成を仮想計算機毎に可能とする仮想計算機システムの入出力制御方法及び装置を提供する。

【0020】本発明の他の目的は、入出力装置からの再結合要求に対し複数のアクセスパスを用いた動的再結合機能を可能とする仮想計算機システムの入出力制御方式を提供する。本発明の他の目的は、リザーブ命令の実行中に障害が起きた場合に他のアクセスパスでリザーブ命令を実行できる拡張リザーブ機能を可能とする計算機システムの入出力制御方式を提供する。

【0021】また本発明の他の目的は、仮想計算機が実行する入出力命令に対するVMモニタの介入を不要にして仮想計算機モニタプログラム(VMモニタ)のオーバーヘッドを削減するようにした仮想計算機システムの入出力制御方式を提供する。本発明の他の目的は、複数の仮想計算機で1つの入出力装置を共用する場合に、入出力制御装置が複数の仮想計算機を認識し、VMモニタの介入を必要とせずに排他制御ができる仮想計算機システムの入出力制御方法及び装置を提供する。

【0022】本発明の他の目的は、複数の仮想計算機で1つの入出力装置を共用する場合にアクセスパスのグループ形成、動的経路再結合機能、および拡張リザーブ機能を可能とする仮想計算機システムの入出力制御方法及び装置を提供する。

【0023】

【課題を解決するための手段】図1は本願の第1発明の原理説明図である。まず第1発明は、仮想計算機システ

ムを制御する仮想計算機モニタプログラム(VMモニタ)22で実現される複数のゲストVM10-1, 10-2と、複数のゲストVM10-1, 10-2の各々に1対1に割り当てられた複数の入出力装置12-1, 12-2と、複数の入出力装置12-1, 12-2の各々に対して少なくとも2系統のアクセスパス(物理アクセスパス)で結合した入出力制御装置(CU)14と、入出力制御装置14を複数のゲストVM10-1, 10-2に結合する少なくとも2つのチャネル装置(CH)16-1, 16-2とを備えた仮想計算機システムの入出力制御方法及び装置を対象とする。

【0024】装置構成を例にとると、第1発明にあっては、入出力制御装置14に、ゲストVM10-1, 10-2からチャネル装置16-1, 16-2及び入出力制御装置14を経由して入出力装置12-1, 12-2に至る複数のアクセスパス111, 112, 221, 222毎に、ゲストVM10-1, 10-2の各々が発行したグループIDを設定したテーブルを保持するグループID保持手段16と、グループID保持手段16のテーブルに設定したグループIDの中の特定のグループIDをクリアするクリア手段18と、ゲストVM10-1, 10-2が発行した入出力命令を受付けた際に、グループID保持手段16のテーブルを参照して入出力装置に至る複数のアクセスパスの間から同一グループに属するアクセスパスを検出し、アクセスパスのグループを形成してゲストVM10-1, 10-2との間の入出力を実行する入出力実行手段20とを設けたことを特徴とする。

【0025】ここで入出力制御装置14に設けたグループID保持手段16は、VMモニタプログラム22による入出力装置12-1, 12-2の割当て結果に基づいてゲストVM10-1, 10-2を起動し、仮想計算機上で動作するオペレーティングシステム(OS)のグループID設定処理によりグループIDの設定を受ける。

【0026】また入出力制御装置14に設けたクリア手段18は、VMモニタプログラム22の指示により特定のグループIDをクリアする。例えばVMモニタプログラム22は、入出力装置12-1, 12-2のゲストVM10-1, 10-2への割り当て変更時に、ゲスト仮想計算機10-1, 10-2が使用していたグループIDをクリアするために入出力制御装置14のクリア手段18にクリアを指示する。

【0027】入出力制御装置14は複数のチャネル装置16-1, 16-2と結合する複数のチャネルアダプタ24-1, 24-2を有し、このチャネルアダプタ24-1, 24-2ごとにグループIDを設定するテーブルを設ける。入出力制御装置14の入出力実行手段20は、保持手段16のテーブルから得られたグループIDで指定された同一のグループに属するアクセスパスを使用して入出力処理を実行する。

10

20

30

40

50

【 0 0 2 8 】更に入出力装置1 2 - 1, 1 2 - 2 は磁気ディスク装置であり、入出力制御装置1 4 は磁気ディスク制御装置である。図2 は本願の第2 発明の原理説明図であり、装置構成を例にとって示している。第2 発明は、複数のゲストVM1 0 - 1, 1 0 - 2 で1 つの入出力装置1 2 - 1 を共用する形態を前提とする。即ち、仮想計算機モニタプログラム2 2 で実現される複数のゲスト仮想計算機1 0 - 1, 1 0 - 2 は、チャンネル装置1 6 - 1、入出力制御装置1 4 を介して入出力装置1 2 - 1 と複数の論理アクセスパス1 1 1 1, 2 1 1 1 と単一の物理アクセスパス1 1 1 で結合される。

【 0 0 2 9 】このようなゲストVM1 0 - 1, 1 0 - 2 による入出力装置1 2 - 1 の共用形態につき第2 発明は、入出力制御装置1 4 に、ゲストVM1 0 - 1, 1 0 - 2 からの入出力命令で指示された論理アクセスパス1 1 1 1, 2 1 1 1 を物理アクセスパス1 1 1 に変換すると共に、ゲストVM1 0 - 1, 1 0 - 2 から入出力装置1 2 - 1 に至る複数の論理アクセスパス1 1 1 1, 2 1 1 1 毎に、ゲストVM1 0 - 1, 1 0 - 2 の各々が設定したグループI Dをテーブルに保持するグループI D保持手段1 6 と、グループI D保持手段1 6 のテーブルに保持した特定のグループI Dをクリアするクリア手段1 8 とゲストVM1 0 - 1, 1 0 - 2 のいずれかが発行した入出力命令を受付けた際に、グループI D保持手段1 6 のテーブル参照により 論理アクセスパスから物理アクセスパス1 1 1 に変換してアクセスの競合を制御する排他制御手段4 0 と、排他制御手段4 0 の制御結果に基づいて入出力装置1 2 - 1 に対する入出力命令を実行する入出力実行手段2 0 とを設けたことを特徴とする。また排他制御以外の点は第1 発明と同様である。

【 0 0 3 0 】

【 作用 】このような構成を備えた本発明の入出力制御方法及び装置によれば、ディスク制御装置( 入出力制御装置) に、磁気ディスク装置( 入出力装置) に対するアクセスパス毎に複数のゲストVMが発行した特定種類のグループI Dを設定できるため、複数のゲストVMの入出力命令に応じてチャンネル装置から与えられるチャンネル・コマンド・ワード( CCW) をディスク制御装置で区別することができ、ゲストVMに磁気ディスク装置を1 対1 に割当てる形態でアクセスパスを2 つ以上とするグループ形成ができる。

【 0 0 3 1 】また複数のゲストVMで単一の磁気ディスク装置を共用する形態にあっても、ディスク制御装置が複数のゲストVMを認識できることで、VMモニタの介在を必要とすることなく排他制御を行うことができ、更にアクセスパスを2 以上とするグループ形成もできる。このグループ形成により、入出力命令により 磁気ディスク装置を起動して突き離し、シーク完了後に再結合する際のアクセスパスの確保が効率良くでき、所謂動的経路再結合機能の実現によって入出力の応答時間を短縮する

ことができる。

【 0 0 3 2 】またゲストVMが磁気ディスク装置をリザーブ中に障害が発生した場合には、別のチャンネルを使用したアクセスができるという所謂拡張リザーブ機能が実現でき、信頼性を向上できる。

【 0 0 3 3 】

【 実施例 】図3 は本発明の入出力制御が適用される仮想計算機システムのハードウェア構成を示した実施例構成図である。図3 において、3 2 は主記憶機装置( MSU) であり、仮想計算機システム上で動作する複数のOSにより実現される複数のゲストVM1 0 - 1, 1 0 - 2、及び仮想計算機システムの監視制御を行うVMモニタ2 2 が設けられている。

【 0 0 3 4 】主記憶装置3 2 には主記憶制御装置( MCU) 2 8 が接続され、主記憶制御装置2 8 を介してストアアクセスまたはロードアクセスを受ける。主記憶制御装置2 8 には複数のCPU3 0 - 1 ~ 3 0 - n が接続されてマルチプロセッサシステムを構成しており、CPU割付けを行うVMモニタ2 2 のディスパッチ機能によりゲストVM1 0 - 1, 1 0 - 2 及びVMモニタ2 2 等のOSを実行する。

【 0 0 3 5 】また主記憶制御装置2 8 に対してはI/Oサブチャンネルとして、チャンネル装置1 6 - 1, 1 6 - 2 を介して1 台のディスク制御装置1 4 が接続され、ディスク制御装置1 4 の配下に複数の磁気ディスク装置1 2 - 1, 1 2 - 2 を2 系統のパスで接続している。ディスク制御装置1 4 にはチャンネル装置1 6 - 1, 1 6 - 2 毎にチャンネルアダプタ( CA) 2 4 - 1, 2 4 - 2 を設けており、チャンネルアダプタ2 4 - 1, 2 4 - 2 はCPUを用いたコントローラ2 6 により入出力制御を受ける。

【 0 0 3 6 】図4 は本願の第1 発明による入出力制御の処理機能を示した説明図であり、OS で実現するゲストVM1 0 - 1, 1 0 - 2 とVMモニタ2 2 に対するチャンネル装置1 6 - 1, 1 6 - 2、ディスク制御装置1 4 及び磁気ディスク装置1 2 - 1, 1 2 - 2 のハードウェア構成と共に示している。ディスク制御装置1 4 は、コントローラ2 6 及びチャンネルアダプタ2 4 - 1, 2 4 - 2 を備え、コントローラ2 6 はグループI D保持部1 6、クリア部1 8 及び入出力実行部2 0 としての機能を備えた一種の処理装置であり、更に、グループI D保持部1 6 により内容が管理されるテーブルを展開する制御記憶2 5 を備える。

【 0 0 3 7 】即ち、グループI D保持部1 6 により、ゲストVM1 0 - 1, 1 0 - 2 からチャンネル装置1 6 - 1, 1 6 - 2 及びディスク制御装置1 4 を経由して磁気ディスク装置1 2 - 1, 1 2 - 2 に至る複数のアクセスパス1 1 1 1, 1 1 2, 2 2 1, 2 2 2 毎に、ゲストVM1 0 - 1, 1 0 - 2 の各々が発行したグループI Dを設定したテーブルを制御記憶2 5 に保持する。

【 0 0 3 8 】またクリア部1 8 により、グループI D保



13

持部16によって制御記憶25のテーブルに格納したグループI Dの中の特定のグループI Dをクリアする。更に入出力実行部20により、ゲストVM10-1, 10-2が発行した入出力命令を受付けた際に、入出力命令を実行する入出力装置が属するグループを制御記憶25のテーブルをグループI D保持部16で参照して同一グループに属する少なくとも2つのアクセスパスを検出し、2つのアクセスパスでグループを形成してゲストVM10-1, 10-2との間の入出力を実行する。

【0039】ディスク制御装置14で実現される機能を更に詳細に説明すると次のようになる。

〔テーブル保持機能〕ディスク制御装置14のグループI D保持部16および制御記憶25によって複数のグループI Dを保持できるテーブル保持機能を実現する。具体的には、磁気ディスク装置12-1, 12-2に対する1つのアクセスパス毎に1つのグループI Dを保持できるようにする。

【0040】例えば図5に示すように、ゲストVM10-1に番号DEV-1の磁気ディスク装置12-1が割り当てられ、またゲストVM10-2に番号DEV-2の磁気ディスク装置12-2が割り当てられていたとする。またゲストVM10-1のグループI D=A、ゲストVM10-2のグループI D=Bとする。この場合、図6に示すように、ディスク制御装置14のチャンネルアダプタ24-1, 24-2の飯盒CA1, CA2で示されるテーブルには、アクセスパス111, 221にグループI D=Aが保持され、アクセスパス112, 222にはグループI D=Bが保持される。

【0041】〔グループ形成機能〕設定グループI Dの中で、同一グループI Dが設定されたアクセスパス同志を同一のグループとみなしてグループ形成する。図6の場合は、ゲストVM10-1に割当てた磁気ディスク装置12-1に対する2つのアクセスパス111, 221のもつ同一グループI D=Aの組で1グループを形成する。またゲストVM10-2に割当てた磁気ディスク装置12-2に対する2つのアクセスパス112, 222のもつ同一グループI D=Bの組で別の1グループを形成する。

【0042】〔グループ設定エラー通知機能〕あるアクセスパスに既にグループI Dが設定済みである時、異なるグループI Dの設定要求に対してエラーを応答する。

〔グループ形成による入出力機能の有効化〕ゲストVM10-1, 10-2からの入出力命令に応じたチャンネル・コマンド・ワード(CCW)を受け付けた際に、同じグループI Dをもつグループ形成されたアクセスパスを認識し、グループ形成により実現する入出力機能を有効化する。このグループ形成により有効化する入出力機能としては、前述したように、①動的経路再結合機能と②拡張リザーブ機能がある。

【0043】〔クリア機能〕特定のグループI Dをクリ

14

アできるグループI Dクリア命令を設け、VMモニタ22がグループI Dクリア命令を発行した時、設定済みの複数のグループI Dの中から命令で指定されたグループI Dを選択してクリアする。このグループI Dクリア命令でグループI Dがクリアされた時は、グループI Dの設定がされていない初期状態に戻る。

【0044】尚、ディスク制御装置14が持つその他の機能は、これらの機能を装備したことによる影響を受けることなく動作する。次に本発明の入出力制御を実現するためにVMモニタ22に対しクリア処理部34の機能が新たに付け加えられる。クリア処理部34は次の条件が得られたときに、チャンネルアダプタ24-1, 24-2ごとに設定済みの特定のグループI Dをクリアする。

【0045】〔条件1〕ゲストVM10-1, 10-2に割り当てていたディスク制御装置14を割当対象から外す場合、具体的にはI/O装置取り付け変更部36に対しオペレータによるゲストVM10-1, 10-2の終了を指示した時或いはゲストVM10-1, 10-2からチャンネル取り外し指示があった時、設定済みの特定のグループI Dをクリアする。

【0046】〔条件2〕仮想計算機システム上のOSが命令エミュレーション部38に対し発行するI/O命令によりチャンネル装置をリセットするクリア指示命令を受けた時、設定済みの特定のグループI Dをクリアする。ここで図4の処理動作を説明すると次のようになる。

【0047】図7はVMモニタ22の処理動作を示したフローチャートである。まず仮想計算機を起動すると、ステップS1で磁気ディスク装置12-1, 12-2をゲストVM10-1, 10-2にVMモニタ22が割り付ける。続いてVMモニタ22からのゲストVM10-1, 10-2の動作開始指示、具体的には、オペレータによってゲストVM10-1, 10-2のIPLが指示され、図8のフローチャートに示す動作が開始される。

【0048】ゲストVM10-1, 10-2の各々は、図8のフローチャートにおけるステップS1のイニシャライズ後のステップS2で、ゲストVM10-1, 10-2の上で動作するオペレーティングシステム(OS)のグループI D設定処理として、チャンネル装置16-1を介してディスク制御装置14に対しグループI D設定命令を発行する。

【0049】このグループI D設定命令は、例えば図9に示すグループI Dをチェーンした形式をもつ。続いてゲストVM10-1又は10-2上で動作するOSが磁気ディスク装置12-1又は12-2に対する入出力要求をステップS3で判別すると、ステップS4で入出力命令をチャンネル装置16-1又は16-2を介してディスク制御装置14に発行し、ステップS4でディスク装置12-1又は12-2のリードまたはライトの完了応答を待つ一連の入出力要求を実行する。この処理は運用停止等のシステム停止がステップS6で判別されるま

50

で継続される。

【 0 0 5 0 】このようなゲスト VM1 0 - 1 又は 1 0 - 2 の入出力命令に VM モニタ 2 2 は一切介在せず、グループ I D で指定されたアクセスパス 1 1 1 , 2 1 1 を使用した磁気ディスク装置 1 2 - 1 に対する入出力処理を実行させる。一方、図 7 に示した VM モニタは、ゲスト VM に対する磁気ディスク装置の割付けを指示した後、ステップ S 2 でオペレータによるゲスト VM1 0 - 1 , 1 0 - 2 の終了指示の有無をチェックし、またステップ S 3 でチャンネルの取り外し指示の有無をチェックし、更にステップ S 4 で OS が発行するチャンネルのリセット命令の有無をチェックしており、いずれかを判別するとステップ S 5 に進む。

【 0 0 5 1 】ステップ S 5 にあつては、VM モニタ 2 2 がディスク制御装置 1 4 のチャンネルアダプタ 2 4 - 1 , 2 4 - 2 ごとに保持したテーブルから該当するゲスト VM が設定したグループ I D を読み込み、そのグループ I D を指定してグループ I D クリア命令をディスク制御装置 1 4 に発行する。このグループ I D クリア命令は、図 1 0 に示すようにグループ I D をチェーンした形式をもつ。このグループ I D クリア命令を受けてディスク制御装置 1 4 のクリア部 1 8 は、VM モニタ 2 2 が指定したグループ I D をテーブルからクリアする。

【 0 0 5 2 】図 1 1 は本願の第 2 発明による入出力制御の処理機能を示した説明図であり、この第 2 発明にあつては 1 台の磁気ディスク装置を複数のゲスト VM で共用する入出力形態を対象とする。第 2 発明の入出力制御を適用する仮想計算機システムのハードウェアは図 3 と同じであることから、図 1 1 にあつては、図 3 のハードウェア構成に対応して各部の機能を示している。

【 0 0 5 3 】図 1 1 において、ゲスト VM1 0 - 1 , 1 0 - 2 は VM モニタ 2 2 により制御される複数のオペレーティングシステムで実現されており、ゲスト VM1 0 - 1 にはグループ I D = A が割り当てられている。また、ゲスト VM1 0 - 2 にはグループ I D = B が割り当てられている。ゲスト VM1 0 - 1 , 1 0 - 2 はチャンネル装置 1 6 - 1 , 1 6 - 2 を通る 2 つの物理的アクセスパスを介してディスク制御装置 1 4 に結合される。

【 0 0 5 4 】ディスク制御装置 1 4 には 2 つのアクセスパスに対応してチャンネルアダプタ 2 4 - 1 , 2 4 - 2 が設けられ、チャンネルアダプタ 2 4 - 1 , 2 4 - 2 を介して、磁気ディスク装置 1 2 - 1 を 2 つの物理アクセスパスで接続している。ここで、磁気ディスク装置 1 2 - 1 はゲスト VM1 0 - 1 , 1 0 - 2 の両方に割り付けられており、ゲスト VM1 0 - 1 , 1 0 - 2 で磁気ディスク装置 1 2 - 1 を共用する。このため、磁気ディスク装置 1 2 - 1 に対するチャンネル装置 1 6 - 1 , チャンネルアダプタ 2 4 - 1 を通る物理アクセスパス 1 1 1 について、ゲスト VM1 0 - 1 , 1 0 - 2 は論理アクセスパス 1 1 1 , 2 1 1 1 を用いて入出力処理を実行する。

【 0 0 5 5 】ディスク制御装置に設けたコントローラ 2 6 にはプログラム制御により実現される保持部 1 6 , クリア部 1 8 , 入出力実行部 2 0 の機能に加え、第 2 発明にあつては更に磁気ディスク装置 1 2 - 1 を 2 つのゲスト VM1 0 - 1 , 1 0 - 2 で共用する際の排他制御を行う排他制御部 4 0 を設けている。また、コントローラ 2 6 の制御記憶 2 5 にはゲスト VM1 0 - 1 , 1 0 - 2 からの論理アクセスパス 1 1 1 1 , 2 1 1 1 を使用した入出力要求に対し、1 台の磁気ディスク装置 1 2 - 1 に対応する物理アクセスパス 1 1 1 に変換するテーブルを備えている。

【 0 0 5 6 】更に、制御記憶 2 5 に展開されるテーブルには、グループ I D 保持部 1 6 にゲスト VM1 0 - 1 , 1 0 - 2 の各々が発行したグループ I D 設定命令に基づいてグループ I D が設定される。このグループ I D の設定は、テーブルに設けられた論理アクセスパス 1 1 1 1 , 2 1 1 1 毎に保持される。

【 0 0 5 7 】更に図 1 1 にあつては、磁気ディスク装置 1 2 - 1 に対するアクセスパスとしてチャンネル装置 1 6 - 2 , チャンネルアダプタ 2 4 - 2 を通る他の物理アクセスパス 2 2 1 を設けており、グループ形成による拡張リザーブ機能や動的再結合機能を実現するため、物理アクセスパス 2 2 1 側についてもゲスト VM1 0 - 1 , 1 0 - 2 からの論理アクセスパス 1 2 2 1 , 2 2 2 1 を割り付け、論理アクセスパスから物理アクセスパスへの変換及び論理アクセスパス毎のグループ I D の設定を物理アクセスパス 1 1 1 側と同様に制御記憶 2 5 のテーブルに行っている。

【 0 0 5 8 】図 1 2 は図 1 1 における仮想計算機 VM1 0 - 1 に対する磁気ディスク装置 1 2 - 1 の割付けを示したもので、番号 DEV - 1 で示される磁気ディスク装置 1 2 - 1 はゲスト VM1 0 - 1 , 1 0 - 2 の両方に割り付けられることで共用される。図 1 3 はコントローラ 2 6 の制御記憶 2 5 に展開されたテーブルの説明図であり、ディスク制御装置 1 4 のチャンネルアダプタ 2 4 - 1 , 2 4 - 2 で決まるチャンネルアダプタ番号 CA 1 , CA 2 をインデックスとして物理アクセスパス 1 1 1 , 2 2 1 を格納している。

【 0 0 5 9 】第 2 発明は、基本的には 1 つの物理アクセスパスを使用して複数のゲスト VM で共用することを前提とすることから、例えばチャンネルアダプタ番号 CA 1 の物理アクセスパス 1 1 1 を例にとると、この物理アクセスパス 1 1 1 を使用して磁気ディスク装置 1 2 - 1 をアクセスするゲスト VM1 0 - 1 , 1 0 - 2 の論理アクセスパス 1 1 1 1 , 2 1 1 1 の 2 つが格納され、論理アクセスパスを参照することで、対応する物理アクセスパスを知ることができる。

【 0 0 6 0 】また、論理アクセスパス 1 1 1 1 については、グループ I D = A が設定され、ゲスト VM1 0 - 1 に属するアクセスパスであることを示している。また、

論理アクセスパス2 1 1 1 にはグループI D=Bが設定され、ゲストVM1 0 - 2 に属する論理アクセスパスであることを示している。グループI Dによるグループ形成は、チャンネル番号C A 2 で示される物理アクセスパス2 2 1 に対応するゲストVM1 0 - 1 , 1 0 - 2 からの論理アクセスパス1 2 2 1 , 2 2 2 1 について、それぞれグループI D=A, Bを保持することで実現される。このようなグループI Dを保持したテーブルを参照することで、同じグループI Dをもつ論理アクセスパス1 1 1 1 と1 2 2 1 が同一グループを形成し、また論理アクセスパス2 1 1 1 と2 2 2 1 が別の同一グループを形成していることを認識できる。

【0061】このグループ形成により第1 発明と同様、リザーブ命令を受けた時に、同じグループI Dをもつアクセスパス同士で入出力装置をリザーブする拡張リザーブ機能が実現できる。また、シーク命令を発行して磁気ディスク装置1 2 - 1 を切り離した後に再結合要求を受けた場合、元の物理アクセスパスがビジー状態にあるときには同一グループに含まれる他の論理アクセスパスに対応する物理アクセスパスを認識して動的経路再結合機能を実現することができる。

【0062】更に図1 3 に示すテーブルにあつては、1 つの物理アクセスパスに対し2 つのゲストVMが論理アクセスパスを使用して入出力を行うことから、入出力装置の獲得状態を示す占有フラグを設けており、占有フラグが1 にセットされているとき入出力装置は占有状態にあり、占有フラグが0 にリセットされていれば入出力装置は占有されていないことを示す。リザーブ命令を受けた時、同じグループI Dをもつ論理アクセスパスの占有フラグが1 にセットされ、リリース命令により同じグループI Dをもつ論理アクセスパスの占有フラグが0 にリセットされる。

【0063】図1 1 のコントローラ2 6 に設けた排他制御部4 0 は、占有フラグのセット、リセットをチェックして入出力装置の排他制御を実現する。尚、図1 1 の第2 発明にあつては、単一の磁気ディスク装置1 2 - 1 を2 つのゲストVM1 0 - 1 , 1 0 - 2 で共有する場合を例にとっているが、磁気ディスク装置が複数であつてもよい。

【0064】また図2 のハードウェア構成にあつては、1 台のディスク制御装置1 4 に2 台のチャンネル装置1 6 - 1 , 1 6 - 2 を組み合わせた入出力構成の形態を例にとるものであったが、ディスク制御装置及びチャンネル装置の数はサブシステムの規模に応じて適宜に定められる。また上記の実施例では入出力装置として磁気ディスク装置を例にとるものであったが、これに限定されずに、光ディスク装置、磁気テープ装置等の入出力機器を含む。

【0065】またマルチプロセッサシステムを例にとっているが、シングルフロセッサシステムでもよし、更に、

VMモニタで実現されるゲストVMの数も任意である。

【0066】

【発明の効果】以上説明したように本発明によれば、入出力制御装置の持つ機能を複数のゲストVMに提供することにより、ゲストVMの入出力実行に関する性能の向上と信頼性の向上を図ることができる。またVMモニタによる入出力命令に対する介入処理が不要となり、VMモニタのオーバーヘッドを低減してシステム性能を向上できる。

【図面の簡単な説明】

【図1】第1 発明の原理説明図

【図2】第2 発明の原理説明図

【図3】本発明のハードウェア構成を示した実施例構成図

【図4】第1 発明による入出力制御の処理機能を示した説明図

【図5】第1 発明のゲストVMに対する磁気ディスク装置の割付け説明図

【図6】第1 発明のディスク制御装置に保持されるアクセスパスとグループI Dの対応を示すテーブルの説明図

【図7】入出力に関する図4 のVMモニタの処理動作を示したフローチャート

【図8】図4 のゲストVMの入出力処理を示したフローチャート

【図9】ゲストVMが発行するグループI D設定命令のコマンド形式の説明図

【図1 0】VMモニタに基づいて発行するグループI Dクリア命令のコマンド形式の説明図

【図1 1】第2 発明による入出力制御の処理機能を示した説明図

【図1 2】第2 発明のゲストVMに対する磁気ディスク装置の割付け説明図

【図1 3】図1 1 のディスク制御装置に保持されるアクセスパスとグループI Dの対応を示すテーブルの説明図

【図1 4】ゲストVMに磁気ディスク装置を1 対1 に割付けた従来の入出力制御の説明図

【図1 5】複数のゲストVMで1 つの磁気ディスク装置を共用する従来の入出力制御の説明図

【符号の説明】

1 0 - 1 , 1 0 - 2 : ゲスト仮想計算機( ゲストVM)

1 2 - 1 , 1 2 - 2 : 入出力装置( 磁気ディスク装置)

1 4 : 入出力制御装置( ディスク制御装置)

1 6 - 1 , 1 6 - 2 : チャンネル装置

1 6 : 保持部( 保持手段)

1 8 : クリア部( クリア手段)

2 0 : 入出力実行手段

2 2 : 仮想計算機モニタプログラム( VMモニタ)

2 4 - 1 , 2 4 - 2 : チャンネルアダプタ

2 5 : 制御記憶

2 6 : コントローラ

28 : 主記憶制御装置( M C U )

36 : I O 装置の割 付 け 変 更 部

30 - 1 ~ 30 - n : C P U

38 : 命 令 エ ミ ュ レ ー シ ョ ン 部

32 : 主記憶装置( M S U )

40 : 排他制御部

34 : グループ I D の ク リ ア 処 理 部

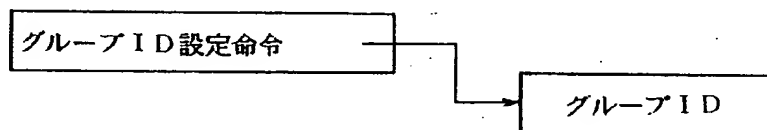
【 図 5 】

## V M モ ニ タ に よ る 磁 気 デ ィ ス ク 装 置 の 割 付 け 説 明 図

仮想計算機	割当磁気ディスク装置
VM10-1	DEV-1
VM10-2	DEV-2

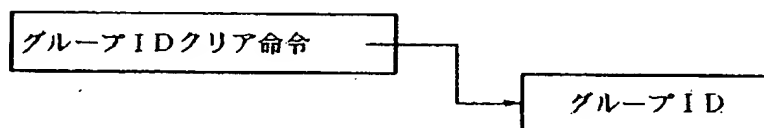
【 図 9 】

## ゲスト V M が 発 行 す る グ ル ー プ I D 設 定 命 令 の コ マ ン ド 形 式 の 説 明 図



【 図 1 0 】

## V M モ ニ タ に 基 づ い て 発 行 す る グ ル ー プ I D ク リ ア 命 令 の コ マ ン ド 形 式 の 説 明 図



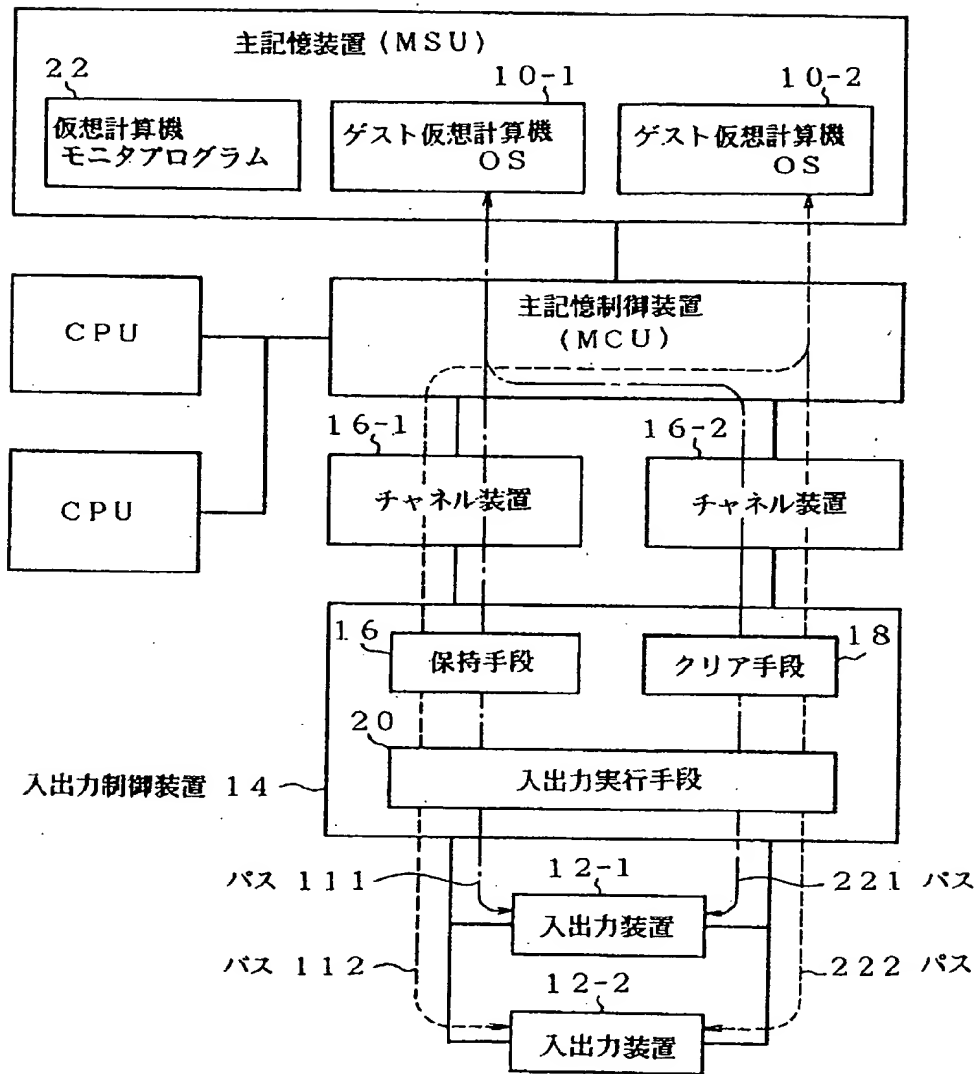
【 図 1 2 】

## 第 2 発 明 の ゲ ス ト V M に 対 す る 磁 気 デ ィ ス ク 装 置 の 割 付 け 説 明 図

仮想計算機	割当ディスク装置
ゲストVM10-1	DEV-1
ゲストVM10-2	DEV-1

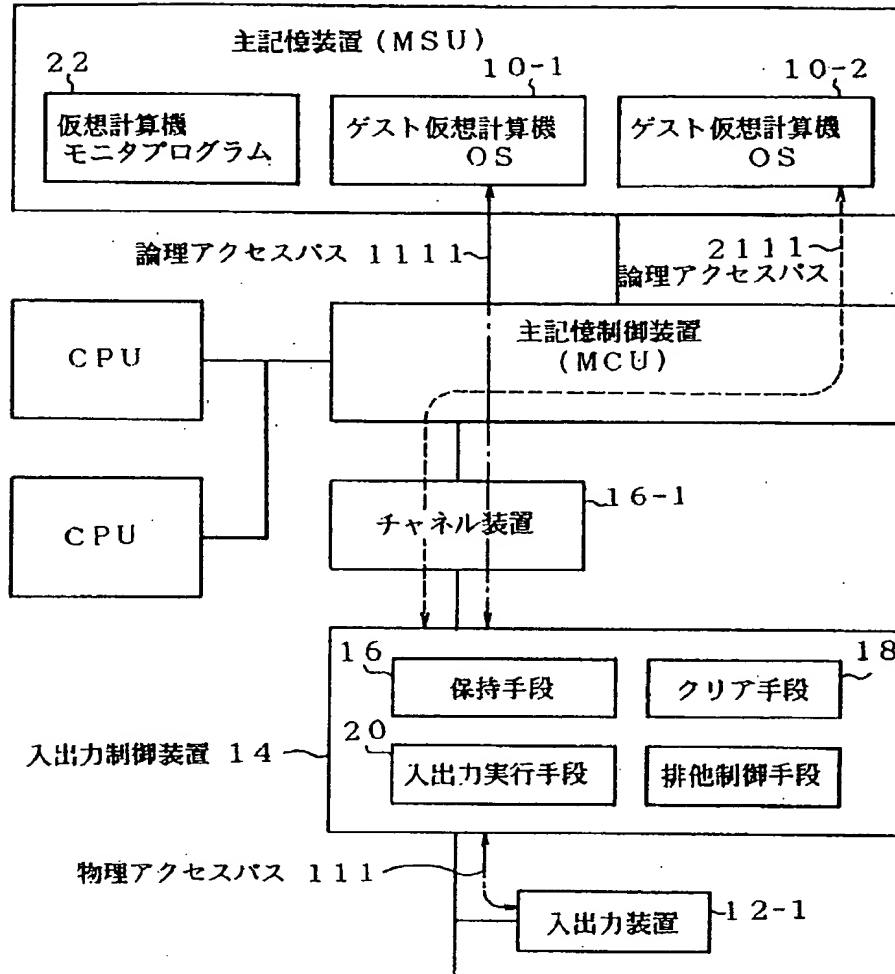
【 図1 】

## 第1 発明の原理説明図



【 図2 】

## 第2発明の原理説明図



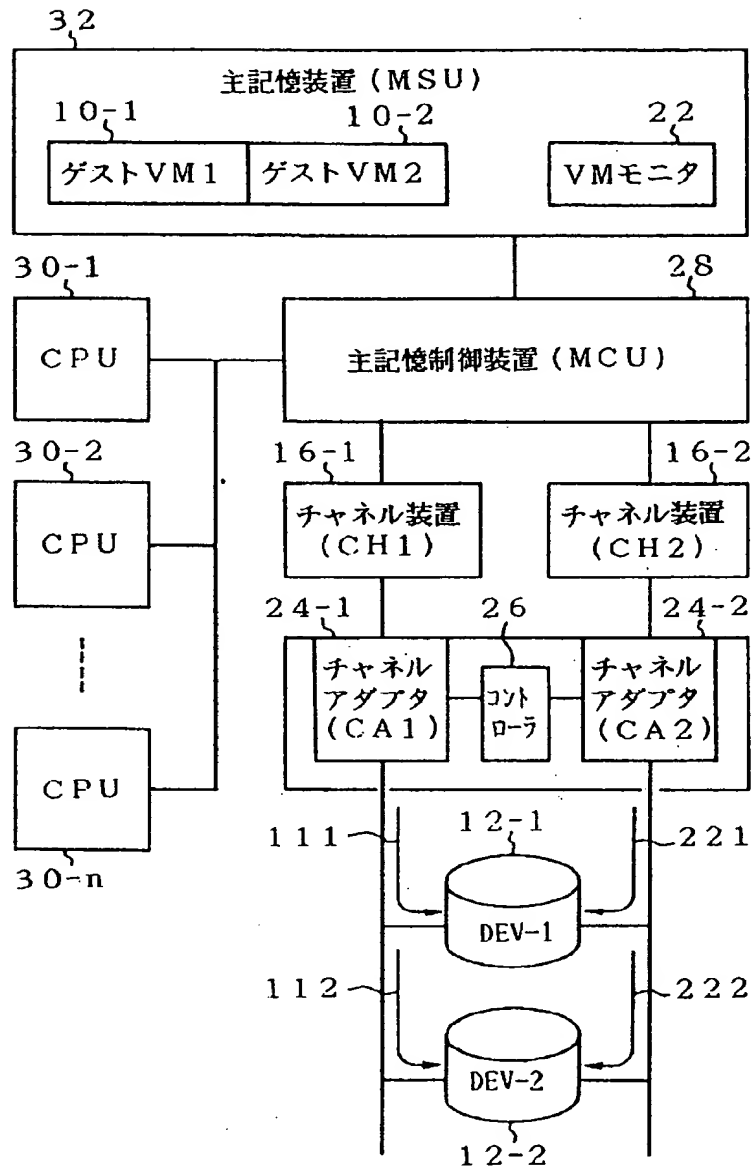
【 図13 】

図11のディスク制御装置に保持されるアクセスバスとグループIDの対応を示すテーブルの説明図

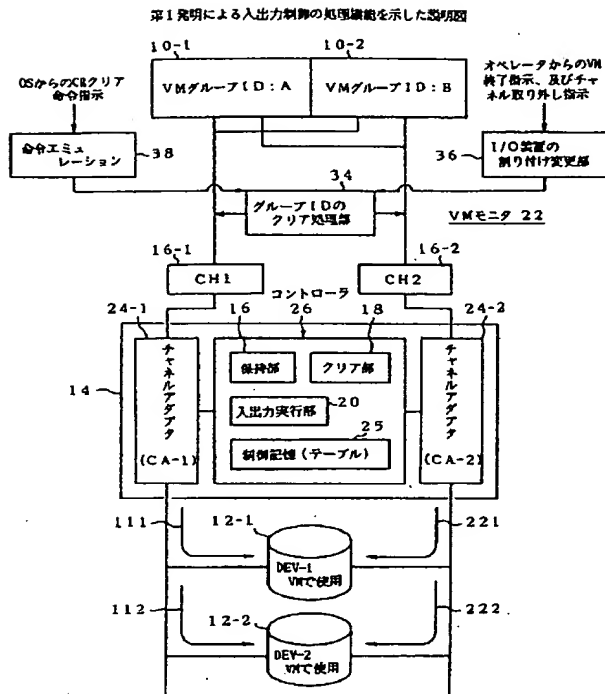
チャネル番号	物理アクセスバス	物理アクセスバス	グループID	格納フラグ	
CA1	111	1111	A	1	グループ形成
		2111	B	0	
CA2	221	1221	A	1	グループ形成
		2221	B	0	

【 図3 】

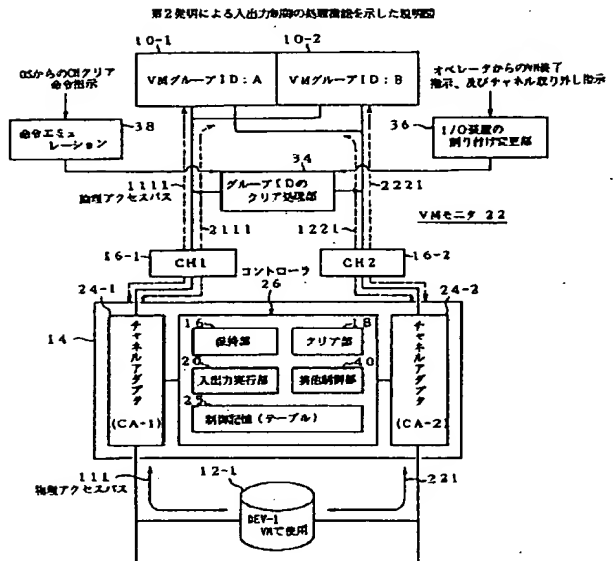
本発明のハードウェア構成を示した実施例構成図



【 図4 】



【 図11 】





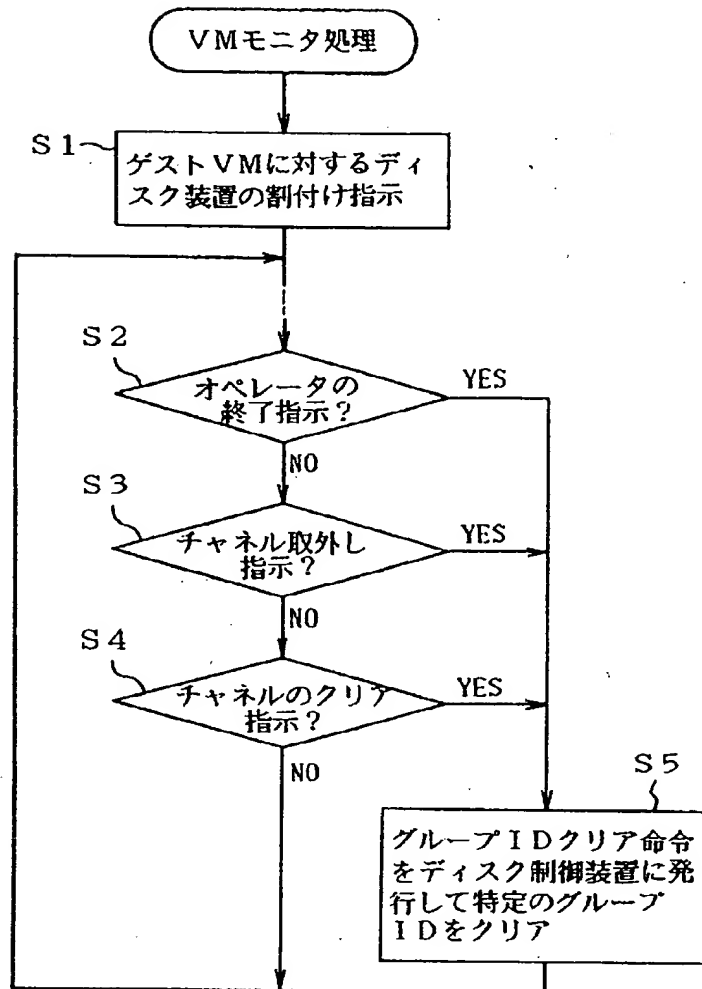
【 図6 】

第1発明のディスク制御装置に保持されるアクセスパスと  
グループIDの対応を示すテーブル説明図

チャンネル アダプタ	パス	設定グループID		
CA1	111	A	グループ形成	グループ形成
	112	B		
	・	・		
	・	・		
CA2	221	A	グループ形成	グループ形成
	222	B		
	・	・		
	・	・		

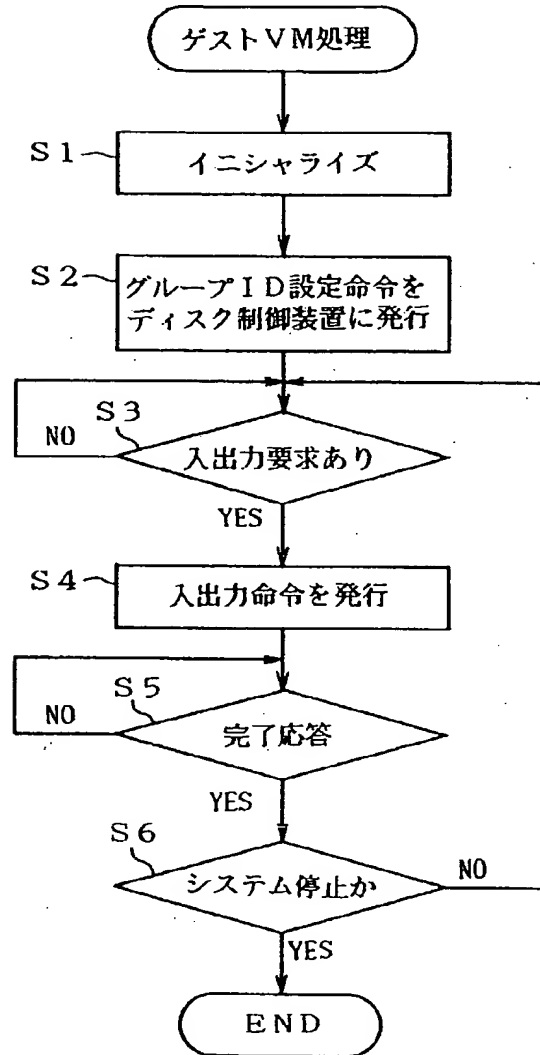
【 図7 】

入出力に関する図4のVMモニタの処理動作を示したフローチャート



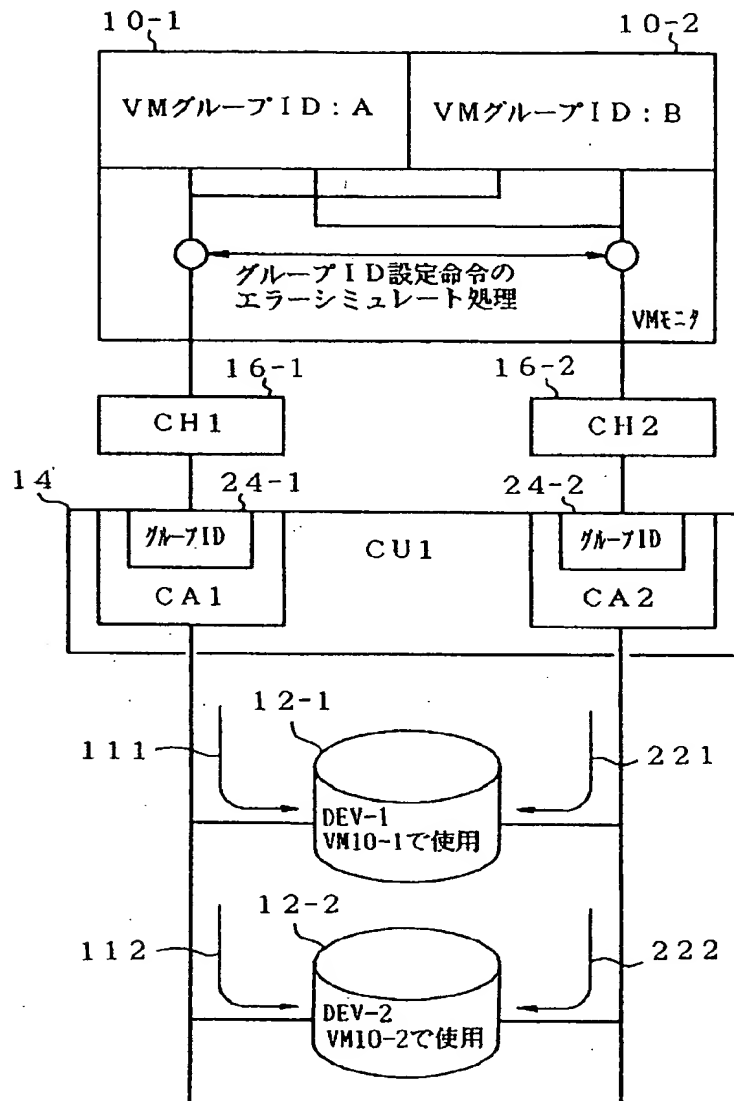
【 図8 】

図4のゲストVMの入出力処理を示したフローチャート



【 図14 】

ゲストVMに磁気ディスク装置を1対1に割付けた従来の入出力制御の説明図



【 図1 5 】

複数のゲストVMで1つの磁気ディスク装置を共用する従来の入出力制御の説明

図

